



Mutational signatures mark cancer's smoking gun

November 3, 2016

LOS ALAMOS, N.M., Nov. 3, 2016—A broad computational study of cancer genome sequences identifies telltale mutational signatures associated with smoking tobacco and demonstrates, for the first time, that smoking increases cancer risk by causing somatic mutations in tissues directly and indirectly exposed to tobacco smoke. The international study by Los Alamos National Laboratory with the UK's Wellcome Trust Sanger Institute and other collaborators was published in the November 4 issue of *Science*.

"This study offers fresh insights into how tobacco smoke causes cancer," said Dr. Ludmil Alexandrov, Oppenheimer Fellow at Los Alamos National Laboratory and co-lead author of the study. "Our analysis demonstrates that tobacco smoking causes mutations that lead to cancer by multiple distinct mechanisms. Tobacco smoking damages DNA in organs directly exposed to smoke as well as speeds up a mutational cellular clock in organs that are both directly and indirectly exposed to smoke."

Tobacco smoke is a complex witch's brew containing more than 7,000 chemicals including over 70 known to cause cancer (carcinogens). Previous large-scale epidemiological studies have associated tobacco smoking with increased risk for 17 different types of cancer, including cancer in tissue not directly exposed to smoke. However, the mechanisms by which tobacco smoking causes cancer have previously remained elusive. This study demonstrates that smoking increases cancer risk by causing somatic mutations that both directly damage DNA and increase the speed of an endogenous molecular clock.

"This research brings together Big Data generated by international cancer consortia and the supercomputing and machine-learning capabilities of Los Alamos to address one of the leading public health issues of our time," said Laboratory Director Charlie McMillan. "Alexandrov's work leverages pattern-recognition software in genomic screening and represents a creative breakthrough in cancer research."

The study, "Mutational signatures associated with tobacco smoking in human cancer," focused on identifying the mutation signatures and methylation changes in 5,243 genome sequences of smoking-related cancers by comparing the cancers of smokers to those of non-smokers.

Professor Sir Mike Stratton, joint lead author from the Wellcome Trust Sanger Institute, Cambridge, UK, said: "The genome of every cancer provides a kind of 'archaeological record,' written in the DNA code itself, of the exposures that caused the mutations that lead to the cancer. Our research indicates that the way tobacco smoking causes cancer is more complex than we thought. Indeed, we do not fully understand the underlying

causes of many types of cancer and there are other known causes, such as obesity, about which we understand little of the underlying mechanism. This study of smoking tells us that looking in the DNA of cancers can provide provocative new clues to how cancers develop and thus, potentially, how they can be prevented."

All mutations--harmless or cancerous--are due to the activity of endogenous or exogenous mutation processes. Each process leaves a signature of scrambled DNA code on the base pairs of that cell's genome. The four base pairs are made up of molecular units identified by the letters A, T, G, and C. The new study found more than 20 mutational signatures across the 17 cancer types associated with tobacco smoking. However, only five of these signatures were elevated in cancers from smokers. Some cancer types had only a single mutational signature elevated in smokers, while others had multiple.

Tobacco smoke and cancer

One signature, called signature 4, can be traced to DNA being damaged by direct exposure to tobacco smoke. "Signature 4 is likely the direct mutational consequence of misreplication of DNA damage induced by tobacco carcinogens," particularly benzo[a]pyrene, according to the study. Signature 5, found by previous Los Alamos research to occur in all cells and to trigger mutations with clock-like regularity, correlated with increased mutations in smokers versus non-smokers. Alexandrov explains that smoking accelerates the clock function mostly likely by altering the molecular machinery underlying this signature.

Other signatures reinforce the theory that smoking increases the risk of several cancer types by raising the overall number of mutations, even in tissue not directly exposed to smoke. The authors note that for some of the mutational signatures, the underlying mechanisms are still unclear.

Alexandrov modeled the cancer mutational processes as a blind-source-separation problem to distinguish coherent signals from a noisy background, a methodology used in other areas of the Laboratory's research related to its nuclear-security mission. The project drew on the Laboratory's high-performance computing resources and expertise, as well as expertise in numerical optimization problems.

Alexandrov is the winner of the 2016 Carcinogenesis Young Investigator Award from the journal *Carcinogenesis: Integrative Cancer Research*, which recognizes a recent significant contribution to carcinogenesis research by an investigator under the age of 40. In 2014, he was recognized by *Forbes* as one of the "30 brightest stars under the age of 30" in the field of science and healthcare. In 2015, he was awarded the *Science* magazine and SciLifeLab Prize for Young Scientists in genomics and proteomics and a Harold M. Weintraub Award for outstanding achievement during graduate studies in the biological sciences. He is a program member of the University of New Mexico Cancer Center.

Researchers representing 16 institutions in the United States, Europe, and Asia worked on the study published by *Science*.

More information, including a copy of the paper, "Mutational signatures associated with tobacco smoking in human cancer," can be found online at the [Science press package](#).

A brief video summarizing the research is available on [YouTube](#).

The [Wellcome Trust Sanger Institute](#)

The Wellcome Trust Sanger Institute is one of the world's leading genome centers. Through its ability to conduct research at scale, it is able to engage in bold and long-term exploratory projects that are designed to influence and empower medical science globally. Institute research findings, generated through its own research programs and through its leading role in international consortia, are being used to develop new diagnostics and treatments for human disease.

Los Alamos National Laboratory

www.lanl.gov

(505) 667-7000

Los Alamos, NM

Managed by Triad National Security, LLC for the U.S Department of Energy's NNSA

